

## PATENT APPLICATION

### METHOD AND CIRCUIT FOR DE-SKEWING DATA IN A COMMUNICATION SYSTEM

Inventor(s): Andy P. Annadurai, a citizen of The United States, residing at  
5112 Amberwood Dr.  
Fremont, CA 94555

Feng Han, a citizen of The United States, residing at  
1541 Poppybank Court  
Pleasanton, CA 94566

Mohammed Rahman, a citizen of The United States, residing at  
3689 Huff Ct.  
Pleasanton, CA 94588

Chris Tsu, a citizen of The United States, residing at  
13625 Surrey Lane  
Saratoga, CA 95070

Assignee: SynTera Communications, Inc.  
42501 Albrae Street  
Suite 203  
Fremont, CA 94538

Entity: Small

## METHOD AND CIRCUIT FOR DE-SKEWING DATA IN A COMMUNICATION SYSTEM

### BACKGROUND OF THE INVENTION

5 [01] The present invention relates generally to the field of communication systems and more specifically to a method and circuit for de-skewing data in communication systems.

10 [02] A data-bus typically comprises a number of parallel interface lines for communicating data signals in parallel. However, the data signals in all of the parallel lines are not always received by a receiver at the same time, even though the data signals are simultaneously transmitted in synchronization with the transmitter system clock. This results in a condition known as data skew. Data skew refers to the delay and/or early arrival of a data bit relative to other data bits in the parallel lines. Conventional systems for de-skewing data on computer data -buses are generally well known. While systems for de-skewing data have been applied to other contexts, such systems are inapplicable to certain communication networks such as SONET (synchronous optical network). In part, this is due to certain unique attributes possessed by SONETs. Another reason relates to certain interface standards that SONETs must meet.

15 [03] SONET is a transport mechanism for multiplexing high-speed data from multiple networks onto fiber optic cabling. It interconnects IP (Internet protocol) networks, ATM (asynchronous transfer mode) switches, T1s and the like via fiber optic networks known as SONET rings. For example, an IP network in San Francisco may be interconnected to a New York IP network via a SONET ring. Since different network types are interconnected, routers are employed for converting protocols from one network type to another. In the case of IP protocols, routers are used for converting the IP packets into a form suitable for transmission on the SONET ring.

20 [04] A board or line card in the router receives the IP packets on one end, performs the necessary conversions and sends out the packets at the other end. Among other components, a line card includes a framer chip for framing/de-framing IP packets, and a system chip for performing protocol conversion and for controlling packet traffic. The framer chip communicates with the system side through an interface known as the SPI-4 (system packet interface). The SPI-4 interface is a standard defined by a consortium of communication companies for packet and cell transfer between a physical layer device (i.e.,

the framer chip) and a link layer device (i.e., the system chip). As defined, the SPI-4 interface has a 16 bit data bus at a clock speed of 622 MHz. Other features of the SPI-4 interface include a transmit control signal (TCTL) and a receive control signal (RCTL) between the system chip and the framer chip. Further, the SPI-4 interface requires the system  
5 chip to periodically send training sequences to the framer chip for detecting data skew. Alternatively, the training sequences are periodically sent from the framer chip to the system chip.

[05] As noted, data skew refers to the delay and/or early arrival of a data bit relative to other data bits in the parallel lines. The skew may be due to a combination of  
10 board layout, environment temperature variation and other reasons. To facilitate understanding the present invention, Figure 1 is a block diagram of a board level connection 100 for the transmit direction between a system chip 100 and a framer chip 101 illustrating data skew on the 16 bit data-bus.

[06] In Fig. 1, the transmit direction has 16 data transmit lines and a transmit control line between system chip 100 and framer chip 101. Each transmit line has a delay buffer 104-1 thorough 104-17, respectively, each representing the total delay caused by board level wires as well as the output delay from system chip 100 and the input delay from framer chip 101. Because the data lines may not be of exactly the same length or same width for all 16 data transmit lines, data bit on each line may each arrive at their destinations 103-1 through 103-17 at different times. Given a clock frequency of 622 MHz (i.e., clock period = 1.6ns) the time differences among the 16 data-bits can be as long as 1 clock cycle or greater thus causing data skew among the data lines. When data skew occurs, all of the data-bits are no longer aligned.  
15  
20

[07] Therefore, there is a need to resolve the aforementioned disadvantage  
25 of data communication networks such as SONET and the present invention meets this need.

#### BRIEF SUMMARY OF THE INVENTION

[08] According to a first aspect of the present invention, a method of de-skewing data in a data communication network such as the synchronous optical network  
30 (SONET) is disclosed. The data is sent from a system chip to a framer chip where the data is de-skewed. To detect data skew, the system chip sends a training sequence to the framer chip. In accordance with the present invention, information bits sent to the framer chip are searched in order to detect the training sequence.

[09] The training sequences contain clear transition patterns at which all of the multiple parallel bits of the transmit data and the TCTL signal line are inverted. If any bit does not invert, this bit is detected as a skewed bit. Based on the value of the data one clock cycle before and one clock cycle after this transition, the skewed bit can be corrected. In one embodiment after the data skew is detected, a multiplexing logic circuitry is used to correct the skew based on one clock cycle either before or after the transition.

[10] According to another aspect of the present invention, a method of de-skewing data in a SONET is taught. The SONET has a system chip for communicating a plurality of data-bits to a framer chip through a data-bus. Among other acts, the method includes: forwarding a sequence of training bits from the system chip to the framer chip; and thereafter, receiving the sequence of training bits at the framer chip. Next, the sequence of training bits received is compared to the sequence forwarded in order to determine if one training bit has a maximum skew of +/-1 clock cycle. If the training bit is not skewed, a no-skew input is selected for receiving the plurality of data-bits.

[11] On the other hand, if the training bit is skewed, it is determined whether there is a late skew or an early skew. If a late skew exists, the late skew is corrected by selecting a late skew input for receiving the plurality of data-bits. It should be noted that the data-bits at the second input are at least one clock cycle earlier than the data-bits for the no skew input. If there is an early skew, however, the early skew is corrected by selecting an early skew input for receiving the plurality of data-bits. The data-bits at the early skew are at least one clock cycle later than the data-bits at the no skew input.

[12] According to another aspect of the present invention, after determining that a data skew exists, any one of three inputs is selected to receive the plurality of data-bits, wherein a first input is selected if there is no data skew, a second input is selected if there is a late skew, or a third input is selected if an early skew occurs.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[13] Figure 1 is a block diagram of a board level connection between a system chip and a framer chip illustrating data skew on an exemplary 16 bit data-bus.

[14] Figure 1A is an exemplary block diagram of a board level connection between a system chip and a framer chip illustrating both the receive and transmit directions.

[15] Figure 2 is a table illustrating an exemplary training sequence for de-skewing bit arrival times on data and control line in accordance with a first embodiment of the present invention.

[16] Figure 3 is a received training sequence illustrating early data skew, the training sequence being received by the framer chip.

[17] Figure 4 is a training sequence for illustrating late data skew, the training sequence being received by a framer chip.

5 [18] Figure 5 is a training sequence for illustrating both early and late data skew, the training sequence being received by a framer chip.

[19] Figure 6 is a block diagram of a multiplexor circuitry for correcting data skew on a 16 bit data-bus in accordance with an exemplary embodiment of the present invention.

10 [20] Figure 7 shows a state diagram (machine) for detecting a training sequence in accordance with an exemplary embodiment of the present invention.

[21] Figure 8 is a block diagram of an exemplary control circuit for the selection lines of the multiplexor of Figure 6.

15 [22] Figure 9 is a block diagram of an alternate multiplexor circuitry for correcting data skew.

[23] Fig. 10 is a generic circuit for correcting bit arrival time in accordance with the present invention.

20 [24] A further understanding of the nature and advantages of the present invention herein may be gained by reference to the remaining portions of the specification and the attached drawings. Further features and advantages of the present invention, as well as the structure and operation of various embodiments of the present invention, are described in detail below with respect to the accompanying drawings. In the drawings, the same reference numbers indicate identical or functionally similar elements.

25 DETAILED DESCRIPTION OF THE INVENTION

[25] In a communication system such as a synchronous optical network (SONET), a framer chip communicates with a system chip through a system packet interface (SPI) standard such as SPI-4 (10 Gb/s) or SPI-5 (40 Gb/s). The SPI-4, for example, has a 16 bit data-bus in the transmit direction (TDAT[15:0]) where [15:0] indicates bits 0 through 30 15, and has an associated transmit control signal (TCTL). Further, the bit data bus has a 16 bit data-bus in the receive direction (RDAT[15:0]) with a corresponding receive control signal (RCTL). According to a first aspect of the present invention, a method of de-skewing TDAT[15:0] in the SONET is disclosed.

[26] Prior to de-skewing, it is determined if a data skew exists by sending a training sequence from the system chip to the framer chip. If there is no skew, a no-skew input of a multiplexor circuit is selected for receiving the plurality of data-bits. If a skew does exist, it is determined whether there is a late skew or an early skew. If there is a late  
5 skew, it is corrected by selecting a late skew input for receiving the plurality of data-bits. Otherwise, the early skew is corrected by selecting an early skew input of the multiplexor.

[27] Figure 2 is a table illustrating an exemplary training sequence 200 for de-skewing bit arrival times on data lines 105-2 through 105-17 (Fig. 1) and control line 106 in accordance with a first embodiment of the present invention. In the present embodiment,  
10 training sequence 200 is compliant with the OIF-SPI4-01.0 and 02.0, phases 1 and 2 defining an interface between physical and link layer, or peer to peer entities operating at an OC-192 rate, and an interface for physical and link layer devices, respectively.

[28] To de-skew data arriving at an interface, training sequence 200 is sent either by system chip 100 (Fig. 1) or framer chip 100 depending on which interface de-skewing is required. For example, if data de-skew is required on framer chip 101 interface, training sequence 200 is sent by system chip 100 to the framer chip interface. When framer chip 101 receives the training sequence, it compares the received sequence with the training sequence forwarded by system chip 101. If the sequences match up, there is no data skew. Otherwise, data skew is present. Data de-skew will also be needed by system chip 100 in the receive direction. When this is required, training sequence 200 is sent by framer chip 101 in the receive direction as shown in Fig. 1A. In Fig. 1A, both the receive and transmit directions are shown. The training sequence for the receive direction is sent as RDAT[0] through RDAT[15], that is via lines 107-17 through 107-2.  
15  
20

[29] As shown, training sequence 200 comprises one or more repetitions of a known training pattern of 1s and 0s. The training pattern includes training data words TDAT [i] (transmit data) where i is the number of data lines. Training sequence 200 also includes a number of TCTL (transmit control) words and a cycle wherein each word is sent. In this exemplary embodiment, the length of training is at least 21 cycles as shown at 202 wherein the number of cycles depend on  $\alpha$ , a system dependent variable. As noted, if the  
25 data received at the destination is the same as those in Figure 2, there is no skew among the data lines. When there are skews among the 16 bit data lines, there can be three situations  
30 illustrated in Figures 3, 4 and 5.

[30] Figure 3 is a received training sequence 300 illustrating early data skew, the training sequence being received by framer chip 101. In Figure 3, during cycle 11, bit 10 of transmit data (TDAT[10]) is one clock cycle ahead of the other bits. That is, TDAT[10] is 0 where it should be 1. Because TDAT[10] should be 0 during cycle 12 and not cycle 11, it is one cycle early.

[31] Figure 4 is a received training sequence 400 illustrating late data skew, the training sequence being received by framer chip 101. In Fig. 4, in cycle 12, TDAT[11] is 1 clock cycle behind the other bits in the table.

[32] Figure 5 is a received training sequence 500 illustrating both early and late data skew, the training sequence being received by framer chip 101. In Figure 5, TDAT[10] is 1 clock cycle ahead, while TDAT[8] is 1 clock cycle behind the other bits.

[33] The above figures illustrate the basic idea of de-skew. First, a known data sequence is sent through the interface. Next, the data at the destination is compared bit by bit with the original data. For each bit, if the value at the destination matches the value at the source, there is no skew for this bit. If the value at the destination mismatches the value at the source, a data skew has taken place. It should be observed that, in this example, the data at each bit remains the same value for 10 clock cycles and then it is inverted, and remains at the inverted value for another 10 clock cycles, when it is inverted back. Thus every 20 clock cycles makes up a complete training period. It is to be understood that the specific numbers of clock cycles are exemplary.

[34] The total number of periods in the training process is determined by the value  $\alpha$ . The tasks for the de-skew include the following: (1) searching for the training sequences; (2) identifying the transition instances; (3) comparing the received values at the destinations with the source values, if there is a match, there is no skew, otherwise, a skew exists; (4) based on the values of the received data, and TCTL 140, determine if the skew is an early skew or a late skew; and (5) based on the results of (4), correct the skews.

[35] The above five tasks make up the three major portions of the circuit to de-skew the transmit data, namely the search for the training sequences (tasks 1 and 2), the determination of early or late skews for every bit in the data bus (tasks 3 and 4), and the data path to correct any possible skews.

[36] Figure 6 is a block diagram of circuitry 600 according to an exemplary embodiment of the present invention for correcting data skew on a 16 bit data bus such as DataIn[15:0] 604.

[37] Among other components, circuitry 600 comprises a multiplexor 620 and three 16-bit data registers namely a first register 611, a second register 612, and a third register 613. Multiplexor 620 is a 3:1 mux having three inputs and one output. The first input is coupled to an output port DataReg1[15:0] 605 of first register 611. The second input 5 is coupled to an output port DataReg2[15:0] 606 of the second register 612 while the third output is coupled to an output port DataReg3[15:0] 607 of the third register 613. Further, DataReg1[15:0] 605 is connected to the only input of the second register 612 while DataReg1[15:0] 606 is coupled to the input of the third 613.

[38] A 16 bit data bus DataIn[15:0] 604, is registered three times at 10 DataReg1[15:0] 605, DataReg2[15:0] 606, and DataReg3[15:0] 607. The data from the above three registers go through mux 620 with two selection lines, ahead[15:0] 609 and behind[15:0] 610, and the output of the mux 620 is the de-skewed data DataOut[15:0] 608. For each bit on the data bus, if there is no skew, the data on DataReg2[15:0] is selected. If 15 the bit is ahead by 1 bit, the data on DataReg3[15:0] is selected. If the bit is behind by 1 bit, the data on DataReg1[15:0] is selected. The selection lines, ahead[15:0] 609, and behind[15:0] 610, are determined in the control logic portion of the circuit. One of ordinary skill in the art will realize that circuitry 600 is illustrative and may be replaced by other circuits consistent with the spirit and scope of the present invention.

[39] Figure 7 shows a state diagram (machine) 700 for searching the 20 training sequences in accordance with an exemplary embodiment of the present invention.

[40] State machine 700 initializes in an idle 110 state where there is no training sequence. When the first pattern of the Transmit Control and the Transmit Data {tctl, tdat} = 1\_0000\_1111\_1111\_1111 is detected (as shown in cycle 2 in Figure 2), state diagram 700 enters a 1st8 state 120 where a pattern of eight consecutive patterns of {tctl, 25 tdat} = 1\_0000\_1111\_1111\_1111 in eight clock cycles are anticipated. If during this period of time, any received pattern is not 1\_0000\_1111\_1111\_1111, state machine 700 goes back to idle 110 state. When eight consecutive training sequences are received, there are two possibilities for the next pattern: (1) the 9th pattern is either the training pattern again, or (2) it is not the training pattern any more.

[41] For the first situation, the state machine goes to a 9th 112 state and for 30 the second situation, state machine 700 goes to an E&L1 116 state. 9th 112 state corresponds to either a no skew situation or one of the situations in Figures 2 or 3, where there are either all early skews or all late skews. The E&L1 116 state corresponds to the situation in Figure 4, where the skews are both early and late.

[42] From 9th 112 state, if the next pattern is the training pattern again, that means all 10 patterns received are training patterns, and there is no skew detected. Thus the state machine enters the noskew114 state and from there it goes back to idle 110 state. Another possibility from 9th 112 state is that the 10th pattern is not the training pattern. In 5 this situation, if TCTL 140 is low, the corresponding bit is late by 1 clock cycle, and the state machine goes to late 113 state where it goes back to idle 110 state in the next clock cycle. If TCTL 140 is high, the corresponding bit is early by 1 clock cycle, and the state machine enters an early 115 state from where it goes back to idle 110 state. The situation in Figure 5, where both early and late skews take place, are covered in E&L1 state 116 and E&L2 state 10 117. The state machine goes back to idle 110 state from E&L2 117 state.

[43] Figure 8 is a block diagram of a control circuitry 800 for the selection lines of multiplexor 620 of Figure 6.

[44] Although not shown, control circuitry 800 comprises logic circuitry implemented by known logic techniques based on desired inputs and outputs of the present invention. In Fig. 8, control circuitry 800 receives three inputs namely de-skewFSM[2:0] 802, TCTL 140 and dataReg1[i] 806 where ( $i = 0, 1, \dots, 15$ ). The outputs of control circuitry 800 are selection lines, ahead[i] 808 and behind[i] 810, where ahead[i] corresponds to 609 and behind[i] is 610 in Fig. 6.

[45] Referring to Fig. 7, the logic for control circuitry 800 is obtained using state machine 700. The values of ahead[i] and behind[i] can be changed only when de-skewFSM[2:0] 802 is in one of the following states: noskew114 state, late 113 state, early 5 115 state, E&L1 state 116 state or E&L2 state 117.

[46] The combination of the circuitry 600, state machine 700 and circuitry 800 of Figures 6, 7, and 8 respectively make up the de-skew circuit. As evident in Figure 6, 25 the longest delay in this de-skew data path is from DataIn 604 through the third data register 607 and it takes three clocks for DataIn to get to the third data register. For timing concerns, it is the best to register the output of DataOut, 606. Thus a total of 4 clock cycles delay may occur to the data going through the de-skew logic.

[47] In this fashion, the present invention provides a circuit and method for 30 de-skewing data in a communication system. While the above is a complete description of exemplary specific embodiments of the invention, additional embodiments are also possible. For example, the above discussion can easily be extended to a skew worse than +/-1 bit time.

[48] For the case of a maximum skew of +/-2 bit time, the data path portion of the de-skew circuitry 900 is shown. In Figure 9, five levels, instead of three levels as in

Figure 6, of registers are needed to store the possible early or late data. A generic circuit for correcting bit arrival time is shown in Fig. 10. Depending on the maximum skew time, N levels are employed for correcting the skew where N is the levels of registers. As shown, the number of ahead selection lines is  $(N-1)/2$  while the number for the behind selection lines is 5 also  $(N-1)/2$ . Modifications for the corresponding state machine and selection lines logic can easily be obtained by one of ordinary skill in the art. It should be observed that while the description of the present invention has been with regard to sending a training sequence in the transmit direction, the training sequence can be sent in the receive direction as well. Furthermore, methods and circuitry presented herein are not limited to a specific SONET 10 system, and may be applied to other types of data communication networks using varying interface standards.

[49] Therefore, the above description should not be taken as limiting the scope of the invention, which is defined by the appended claims along with their full scope of equivalents.

15  
14  
13  
12  
11  
10  
9  
8  
7  
6  
5  
4  
3  
2  
1